REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL – 01 JUL 92 TO 31 DEC 95

4. TITLE AND SUBTITLE

GEOMETRIC INVARIANTS IN OBJECT RECOGNITION

5. FUNDING NUMBERS

F49620-92-J-0332
2304/GS 61102F

6. AUTHOR(S)

ISAAC WEISS

AFOSR-TR-96

0370

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

UNIVERSITY OF MARYLAND
2100 LEE BUILDING
COLLEGE PARK, MD 20742

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NM
110 DUNCAN AVE, SUITE B115
BOLLING AFB DC 20332-8080

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

F49620-92-J-0332

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

SEE REPORT FOR ABSTRACT

19960726 051

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED

NSN 7540-01-280-5500

DTIC QUALITY INSPECTED 1

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std 239-18
299-102

Geometric Invariants in Object Recognition
Final Report
Isaac Weiss

Abstract

The project has involved the use of invariants in computer vision and has shown that invariants can greatly improve the efficiency of object recognition. As part of the project, two kinds of invariants have been studied. 1) Geometric invariants, which assume that the geometry of the object is given. The objective here is to use descriptors of the object that are independent of the geometric transformations that complicate object recognition, such as change of the viewpoint from which the object is seen, or small deformations. ii) Physical invariants, which take into account the physical process by which the image is obtained, such as irradiation by visible light, infra-red, radar, sonar etc. We borrow methods from physics to apply invariants similar to energy and momentum of the physical process. In these processes there are usually more unknowns than equations. The invariants help in putting additional constraints on the underdetermined set of equations.

1. Geometric invariants

Geometric invariants are shape descriptors which remain invariant under geometric transformations such as projection or viewpoint change. They are important in object recognition because they enable us to obtain a description of an object which is independent of the viewpoint and other geometric transformations.

In the object recognition process, we have an image of an unknown object, and we want to match it to an object stored in a database. However, since the image depends on the viewpoint from which the unknown image is seen, we cannot perform the match until we find the correct viewpoint. This involves a search in a high dimensional search space, involving all the parameters that determine the viewpoint—a search which is very costly if it can be performed at all.

Invariants eliminate this high complexity search, thus increasing by orders of magnitude the efficiency of the matching process. The number of objects that can be stored and

1

matched is also greatly increased because we do not need to store different views of the same object. The basic idea is to store in the database not the object description itself, but a set of descriptors derived from it which are invariant to the viewpoint. The descriptors can be numbers, curves or other entities. Each object will have a corresponding set of such invariant descriptors stored in the database. Later, when given an image of an unknown object, we will again derive from it a set of invariant descriptors in the same way we did for the stored objects. We will then match the observed set of invariants with the stored ones. A successful match will give a positive identification without looking for the right viewpoint.

Because of the advantages mentioned above, the subject of invariants has recently gained in importance and acceptance in the vision community. Projective (viewpoint) invariants were a very active mathematical subject in the latter half of the 19th century. However, in vision only one very limited projective invariant, the cross ratio [Duda and Hart, 1973], was used until recently.

Much more general viewpoint invariants were first introduced in vision by Weiss [1988]. In that paper we described invariants of point and line sets, of conics and of general curves, and pointed out their usefulness for object recognition.

Of the invariants we described in the above paper, only simple ones were used prior to the present project. They were invariants of very simple objects, such as points, lines and sometimes ellipses. The present project has been much more general and dealt with general curves. Curves are very strong descriptors of objects, and we can use either boundary contours of the objects or contours on the objects themselves for recognizing the objects. They are much richer descriptors of shapes, capable of describing much more complicated objects then lines or ellipses. Thus, viewpoint invariants of general curves are very helpful in object recognition.

Attempts to use curve invariants before this project were very limited and suffered from several problems. An example of such an older method was plotting the curvature of a curve against its arclength, as follows. Given a curve, choose a starting point s_0 on it. Then, for any other point on the curve, calculate: (a) its distance s along the curve from s_0 (i.e. the length of the arc from s to s_0), (b) the curvature κ at s. Than plot

2

the curvature as a function of the arclength, i.e. plot $\kappa(s)$. Since κ and s are invariant to translations and rotations, we obtain a "Euclidean" invariant curve that can be used instead of the original curve for storage and matching.

There are several problems with the above method. First, it is not fully viewpoint invariant. It is only invariant to translations and rotations, whereas a viewpoint change is much more general; it involves projections between planes having general tilt and slant angles. Second, this invariant curve depends on the choice of the initial point s_0, so this starting points becomes an extra parameter that complicates the search and matching. Third, the method is sensitive to occlusion. That is, if part of the original curve is missing, there is no way to calculate the arclength across the missing part and thus no way to plot the curvature against the arclength.

During the project we have developed a method which overcomes all the above difficulties. Instead of the arclength and curvature, we use two invariants I_1, I_2 calculated at each point of the curve. These invariants are calculated at each point according to an algorithm that we have developed. We then plot one of these quantities, (say I_1), against the other (say I_2). That is, for each curve point we mark another point in an "invariant plane", having coordinates I_1, I_2. The collection of all these points with invariant coordinates I_1, I_2 forms an invariant "signature" curve. This curve can be used for recognition and matching instead of the original curve.

These plots of I_1 vs. I_2 form the basis for our object recognition system. Given an unknown object, we extract some contours describing it, than calculate their invariant signature curves as described above, and then match these signature curves to a library of known signature curves. A successful match indicates that the observed object is similar to the one in the library, even though it is seen from a different viewpoint.

The method just described is well understood in theory. However, implementing it in practice has presented us with several problems that we had to overcome during the project.

One problem was calculating reliable derivatives. The invariants I_1, I_2 that we use require the use of derivatives of the curve, and these are notoriously unstable when the curve has to be extracted from a real image. We have developed a method to calculate

3

derivatives robustly. We have studied in detail the sources of errors that corrupt the derivative values and found ways to eliminate them. The basic idea is to fit certain polynomials over a relatively wide neighborhood around a point, and then calculate the derivatives of the polynomial at that point. We have derived a formula characterizing the trade-off between the accuracy of the derivative, the order of the polynomial, and the width of the neighborhood. Based on these concepts, we have designed filters that can be applied to a curve to find its derivatives. This method is of course of a very general applicability in vision, beyond the applicability for invariants.

Another problem involved small deformations of the curves. Images of curves are often slightly different from each other, even if they represent a similar object. There are several sources of such deformations, including lens aberrations and differences between samples of the same kind of object. For example, different human bodies will have somewhat different outline curves but will retain the general shape of a human. The problem is less noticeable in machine made objects. We have discovered that invariants of certain viewpoint transformations, namely affine transformations, are rather insensitive to these deformations. These transformations are a subset of the general viewpoint (projective) transformations, and they are a good approximation of the general viewpoint transformations when the object is far away from a camera. This condition is usually satisfied in practical applications such as tracking an airplane by a camera. Thus we can usually use affine transformations instead of the general projective ones. The insensitivity of their invariants to small deformations has been shown theoretically and demonstrated experimentally. These invariants can thus be called "quasi-invariants" of small deformations.

A third problem involves discontinuities in the curves. Since our curve invariants depend on derivatives, they cannot be calculated at discontinuities. This problem is a subject of a proposed continuation project. For the present project, we have concentrated on smooth natural objects such as apples and pears which have few discontinuities. The proposed project will deal with man-made objects such as airplanes which contain many discontinuities.

We have demonstrated an application of the methods described above to the recognition of natural objects such as fruits. Our system extracts the outline of the fruit,

calculates its invariant signature curve using the methods described above, and matches to a library of signature curves of various fruits. We have demonstrated, for instance, that a the signature curves of all pears are quite similar even if they are viewed from different angles (except from directly above or below). On the other hand, these pear signature curves are different from those of apples or bananas. This enables us to distinguish a pear from other fruits without searching for the correct viewpoint. We are currently developing a system for a similar recognition of man-made objects.

2. Physical Invariants

The invariants described in previous sections are purely geometrical, i.e. they rely on knowledge of the geometry of the shape. In practice, the geometry is not given directly, but has to be inferred from a physical process involving a physical medium such as visible light, infra-red radiation, radar, sonar, etc. This adds additional unknowns besides the geometry, such as light intensity, surface reflectance characteristics, etc. Normally we have more unknowns than equations and the original shape cannot be recovered without some modeling assumptions.

Invariants can play a key role here by eliminating some of the unknowns of the problem. Invariants of physical processes are very well studied in physics and are of great value there in simplifying problems. For example, the law of energy conservation means that energy is an invariant of the physical process. Using this law, we can calculate various physical quantities without knowing all the exact details of the process. If we drop a bomb from a certain height, we know at what velocity it will hit the ground without calculating its trajectory. This is because the bomb's kinetic energy at the ground level is equal to its potential energy at the height of the airplane.

This concept is greatly generalized in mathematical physics. Besides energy, other invariants are momentum, angular momentum etc. Invariants are usually a result of symmetry of the physical process. For example, conservation of energy is a result of symmetry with respect to time: the law of physics are the same whether we measure them now or at some time in the future. Conservation of momentum is a result of symmetry to translation in space: the laws are the same here as in another lab at a distance from here. The

symmetry refers to the laws of physics themselves, not to the shapes of the objects we are experimenting with or to the instruments.

These symmetries provide the modeling assumptions mentioned above. When a situation possesses a symmetry, one can immediately derive a corresponding invariant constraint that helps solve the problem. The basic result that connects symmetries with invariants is Noether's theorem.

It turns out that the mathematical formalism mentioned above has a wide applicability in vision problem and is not necessarily limited to physical processes. Purely geometrical problems such as shape from texture can be cast in a way suitable to be handled by the physical invariants method. We have concentrated on the shape from shading problem, and found that geometrical entities such as the slopes of the surface are analogous to physical momentum and obey a law of momentum conservation.

During the present project we have laid the theoretical foundations for these ideas. In a proposed continuation project will implement the theory in practical applications.

Publications under the Project

3. Journal Publications

- Smoothed Differentiation Filters for Images (with P. Meer), *Journal of Visual Communication and Image Representation*, **3**, (1), 58-72, March 1992.
- Geometrical Invariants and Object Recognition, *International Journal of Computer Vision*, **10**, 207-231, May 1993.
- Noise Resistant Invariants of Curves, *IEEE T-PAMI*, **15**, 943-948, September 1993.
- High Order Differentiation Filters that Work, *IEEE T-PAMI*, **16**, 734-739, July 1994.
- A Convex Polygon is Determined by its Hough Transform, (with A. Rosenfeld), *Pattern Recognition Letters*, **16**, 305–306, October 1994.
- Local Invariants for Recognition, (with E. Rivlin), *IEEE-PAMI*, **17**, 226-238, March 1995.
- Local Projective and Affine Invariants, *Annals of Mathematics and Artificial Intelligence*, **13**, no. 3–4, 203-225, 1995.
- Recognizing Objects Using Deformation Invariants, (with E. Rivlin), *Computer Vision, Graphics and Image Processing: Image Understanding*, accepted.
- Applying Algebraic and Differential Invariants for Logo Recognition, (with D. Doermann and E. Rivlin), *Machine Vision and Applications*, accepted.
- The Geometry of Visual Space. (with Z. Pizlo and A. Rosenfeld), *Computer Vision and Image Understanding*, accepted.

4. Papers under Journal Review

- Analytic Line Fitting in the presence of Random Noise (with N. Netanyahu).
- Physics-like Invariants for Vision.
- Model-based Recognition of 3D Curves from One View.
- Reconstruction on 3D Curves from Uncalibrated Cameras.
- Scale Space Invariants for Recognition (with A. Bruckstein and E. Rivlin).